DOCUMENT RESUME

ED 173 419 TM 009 467

AUTHOR Huntley, Renee M.; And Others
TITLE The Assessment of Rhetorical Proficiency: The Role of Objective Tests and Writing Samples.
PUB DATE [Apr 79]
NOTE 19p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, California, April 9-11, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Freshmen; Composition Skills (Literary); Expository Writing; Higher Education; *Language Proficiency; *Objective Tests; *Performance Tests; Predictive Ability (Testing); Scores; *Test Validity; *Writing Skills
IDENTIFIERS *American College Test English Usage Test; Scholastic Aptitude Test; Test of Standard Written English

ABSTRACT

An investigation of: (1) the relationship between the American College Test (ACT) English Usage Test scores and demonstrated writing proficiency, and (2) the relationships between and among other objective tests and writing proficiency was undertaken, using three samples of college freshmen. Although results indicated a moderately high relationship between objective measures and writing samples, the skills measured by the two types of instruments were somewhat different. Collectively, results suggested that some writing skills, particularly those represented by the holistic, language, and organization writing scores, appeared to have a moderately high relationship to the objective assessments. The audience writing score exhibited a much lower correlation to the objective assessments and a lower interrater reliability than other writing scores. It was suggested that further research be directed toward studying the influence of scribal errors on raters of writing assessment and the selection of specific writing strategies, processes, and features for objective assessment. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173419

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE ASSESSMENT OF RHETORICAL PROFICIENCY: THE ROLE OF OBJECTIVE TESTS AND WRITING SAMPLES

Renee M. Huntley

Cynthia B. Schmeiser

and

Richard J. Stiggins

American College Testing Program

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Renee Huntley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 009467

Paper presented at the Annual Meeting
of the National Council on Measurement
in Education, San Francisco, 1979

THE ASSESSMENT OF RHETORICAL PROFICIENCY: THE ROLE OF OBJECTIVE TESTS AND WRITING SAMPLES

Renee M. Huntley

Cynthia B. Schmeiser

and

Richard J. Stiggins

American College Testing Program

This study examined the validity of objective tests for the assessment of rhetorical proficiency in several post-secondary settings. Both innovative and traditional objective test formats were administered, compared conceptually, and correlated in an attempt to determine the relationship of each to the assessment of writing proficiency. Implications for the use of objective assessments of rhetorical proficiency are examined.

The content and focus of both the secondary and the postsecondary English curriculum have been subjected to intense scrutiny and reevaluation in response to the current public concern over the seeming deficiencies among students in the skills commonly associated with the English classroom. One consequence has been a renewed interest in the assessment of writing proficiency which, in fact, has had a relatively long history as a focus for research; as far back as 1921, the problems of reliability and validity were being confronted (Godshalk, Swineford, and Coffman, 1966). The academic climate that inspired some of the research in the sixties (e.g., Godshalk, et al.) prevails again: educators, particularly English educators, are insisting that writing proficiency be demonstrated directly.

No one would dispute that writing samples developed and scored according to sound measurement principles provide the most ecologically valid measures of writing proficiency. At the very least, focusing the testing spotlight on actual writing samples would, it is believed, encourage increased attention to writing in the curriculum. But the evaluation of actual samples of student writing is a complex task that involves considerable and crucial decision making. The questions that typically arise would include the following: What is the purpose for the assessment? For example, will it roughly sort students in order to predict performance, or will it precisely describe the features of writing for a given rhetorical task in order to identify problems? Which facets of the written product should be examined? Can we account for the writer's control of writing strategies in the pre-writing, transcribing, rethinking, and editing stages of the process? Do we want to measure writing ability in only one mode or in multiple modes? Does

the measurement scheme really permit an accurate description of writing according to the best current research about writing strategies, processes, products, and modes? And, finally, can we resolve the difficulties inherent in establishing standards of acceptable performance?

Attempts to resolve these questions, like those presented by Steele (1979), can be costly and require a high level of content expertise. It is the premise of this paper that, in those instances where cost and available content expertise are factors, objective tests could be used as an indirect, "passive" way to assess some of the features or components of that complicated, integrated process called composing. Since research has only now begun to focus on these components, the most judicious selection for objective test formats of key features that indicate the whole composing process may presently be somewhat problematic.

Nevertheless, the premise that objective tests can be used to approximate language proficiency has been worth investigating both here and elsewhere. One earlier study concluded that objective tests do "an amazingly effective job of ordering students in the same way as a trained group of readers would after reading a sizable sample of their actual writing" (Godshalk, et al., p. 21). The correlations reported there between eight ETS English Composition Test subtest scores and total writing scores ranged from .458 to .707, with the highest correlation between usage and writing scores and the lowest between paragraph organization and writing scores. More recently, data reported on the relationship between an objective usage exam and a 20-minute writing sample cited moderate correlations of approximately .60 (Breland, 1977).

The study presented in this paper was designed to examine this premise further by using several assessments of language proficiency.

Specifically, the purposes of this study were twofold: 1) to investigate the relationship between ACT English Usage Test scores and demonstrated writing proficiency; and 2) to examine the relationships between and among other objective assessment instruments and demonstrated writing proficiency.

Subjects

To examine the relationship between the ACT English Usage Test, other objective assessment instruments, and demonstrated writing proficiency, data from 3 samples were obtained. The first sample comprised 61 college freshmen from 5 selected postsecondary institutions that participated in the ACT College Outcome Measures Project (COMP). Three writing exercises were administered to these freshmen, and the results were correlated with each student's preadmission standard score on the ACT English Usage Test. The second sample comprised 35 college freshman in 2 intact English classes in a major midwestern university; 2 writing samples were administered to them, and the results were correlated with the students' preadmission ACT English Usage Test scores. (Note: In both samples 1 and 2, the time interval between administration of the ACT Test and writing samples ranged from 1 to 2 years.) The third sample comprised 50 college freshmen in a 4-year southeastern college. An abbreviated version (35 items) of the ACT English Usage Test, an experimental assessment of rhetorical proficiency, and 3 writing samples were administered simultaneously to these students. Preadmission scores from the verbal section of the Scholastic Aptitude Test (SAT-Verbal) and the Test of Standard Written English (SAT-TSWE) were also available for this sample.

Instruments

The objective test instruments used in these studies were the ACT English Usage Test, an experimental objective test developed by ACT to assess rhetorical proficiency, the SAT-Verbal, and the SAT Test of Standard Written English.

The ACT English Usage Test measures the student's *understanding* of (1) the conventions of standard written English and (2) the use of basic elements of effective expository writing (the kind of writing that characterizes the educational environment). The test's format consists of seven prose passages that have underlined and numbered portions. For each underlined portion, four alternative responses are given, one of which is the response that appears in the text. The student is asked to decide which alternative best expresses the idea, makes the statement appropriate for standard written English, or is worded most consistently with the style and tone of the passage as a whole. The 75 items in the test cover 5 elements of effective expository writing: punctuation, grammar, sentence structure, diction and style, and logic and organization.

For data samples 1 and 2, ACT standard scores (1 to 36) were used to represent English Usage scores. In data sample 3, a 35-item abbreviated version of the test was used. In this case, scores were reported on a 35-point raw score scale.

The experimental ACT instruments were designed to assess students' critical perception of such language activities as the author's purpose for communicating, the author's intention, the occasion, and the audience for whom the communication was intended. Passages with underlined segments were presented not for analysis of "error" in usage and mechanics, but to demonstrate the range of discourse appropriate to varying situations,

the shifting of language to suit the occasion, and the effectiveness and honesty of various language options. The items tested (1) the student's ability to make inferences about the author's purposes and motives, the author's underlying assumptions, and the effects created by language choices, and (2) the students' ability to perceive whether the strategies of organization were effective for the purposes of the given piece of discourse (appropriateness of beginning, middle, and ending; options in development; forms of support; and overall unity). These items were constructed to simulate the kind of decision making that confronts a writer. Two different formats were developed: a single-passage unit that contained 18 items (ACT Rhetorical Proficiency Test-Format I); and a unit of 4 sets of paired paragraphs that contained 20 items (ACT Rhetorical Proficiency Test-Format II). Raw score scales were used to represent performance in each case.

The SAT-Verbal includes antonym items, analogies, and reading comprehension. Standard scores (200 to 800) represented student performance on this test. The SAT-TSWE measures "a student's ability to use language in which most college textbooks are written" (CEEB, 1975, p. 5). Areas assessed include grammar, sentence structure, diction, and style. Two item formats are used, each of which asks the student to identify and correct errors in single sentences. Standard scores ranging from 20 to 60 represented student performance on this test.

The 3 writing exercises described by Steele (1979) served as the direct assessment of writing proficiency in the 3 studies. These COMP writing exercises were administered to subjects and rated according to the procedures outlined by Steele. The rating scale, which ranged from 0 (low) to 4 (high), was used to assign 1 holistic score (global judgment) and 3 analytic subscores--*language* (practical concerns of expression),

organization (rhetorical concerns), and *audience* (attending to the perspective of the audience). Each student received 4 scores on each exercise, and the scores were summarized across the raters.

Analysis

To investigate the relationship between the objective assessments and demonstrated writing proficiency, Pearson product-moment correlation coefficients were computed between the scores students' obtained on the objective assessment instruments and their scores on actual writing exercises. Correlations were computed for each of the 3 data samples, separately. In addition, estimates of internal consistency reliability (KR_{20}) were computed for the objective assessment instruments, and generalizability coefficients reflecting interrater agreement were computed as estimates of the reliability of the essay ratings.

RESULTS

For each of the two purposes of this study, the results are presented below.

First, what is the relationship between the ACT English Usage Test and demonstrated writing proficiency?

The results of the analysis of the relationship between the ACT English Usage Test and demonstrated writing proficiency are reported in Table 1 for data samples 1 and 2, and in Table 2, for data sample 3.

The correlations based on data samples 1 and 2 suggest that a moderate correlation exists between the preadmission ACT English Usage Test scores and holistic ratings of writing performance. The correlations based on data sample 3, being somewhat higher than those for data

TABLE 1

Correlations Between High School
ACT English Usage Performance and
Writing Performance Measured in College

SAMPLE	N	WRITING PERFORMANCE			ACT		CORRELATION
		EXERCISES	MEAN	S.D.	MEAN	S.D.	
1	61	3	1.93	.64	21.46	5.22	.43
2	35	2	1.74	.85	21.85	2.72	.50

TABLE 2

Correlations Between Objective Tests and Measures of Writing Proficiency

(N = 50)

	WRITING SCORE						
OBJECTIVE TESTS	HOLISTIC	LANGUAGE	ORGANIZATION	AUDIENCE	INTERNAL CONSISTENCY RELIABILITY	MEAN	S.D.
CT ENGLISH USAGE (35 items)	.67	.71	.69	.44	.83	22.30	6.12
ACT RHETORICAL PROFICIENCY-FORMAT I (18 items)	.51	.43	.61	.25	.63	9.62	3.25
ACT RHETORICAL PROFICIENCY-FORMAT II (20 items)	.53	.48	.58	.31	.73	10.44	3.81
ACT TOTAL (ENGLISH USAGE, RHETORICAL PROFICIENCY-FORMATS I AND II) (73 items)	.69	.68	.75	.44	.89	42.16	11.15
SAT-VERBAL	.58	.56	.60	.38	.90	359.40	133.42
SAT-TSWE	.66	.69	.71	.38	.85	39.92	11.93
INTERRATER RELIABILITY (2 raters)	.77	.86	.74	.48			
MEAN*	10.52	12.64	10.16	11.30			
S.D.	4.97	4.37	4.71	4.97			

*Sum of 6 ratings on 0 to 4 scale (3 exercises, 2 raters)

samples 1 and 2, suggest that the relationship between postadmission ACT English Usage Test scores and writing proficiency scores (holistic, *language*, and *organization*) is moderately high. Perhaps the difference in the magnitude of the correlations for the 3 data samples can be attributed to the time of administration of the English Usage Test (preadmission or postadmission). However, the correlation between the English Usage Test scores and the *audience* subscore of the writing sample is low to moderate.

Second, what is the relationship between and among other objective assessment instruments and demonstrated writing proficiency?

Table 2 presents the results of the analysis of the relationship between all the objective assessment instruments and demonstrated writing proficiency, based on data sample 3. In general, the relationships between the objective assessments and the holistic, analytic-*langua* and analytic-*organization* writing subscores are moderately high. The correlations between the objective scores and the *audience* writing subscore are, once again, low to moderate.

In addition, the intercorrelations among the various objective tests that were computed are presented in Table 3. Examination of these intercorrelations suggests that the measures seem to correlate moderately highly with each other. The lowest correlations occur between the two formats of the ACT measure of rhetorical proficiency and the other objective (predominantly usage) measures; the skills measured on the ACT rhetorical proficiency instruments would, therefore, seem to differ from those measured in the other objective instruments.

As reported in Table 2, the reliabilities of the objective instruments used in this study are all moderate to high for generally short tests. This would indicate that these measures generally are internally consistent.

TABLE 3

Intercorrelations Among
Objective Tests

		1	2	3	4	5	6
1.	ACT ENGLISH USAGE	-					
2.	ACT RHETORICAL PROFICIENCY-FORMAT I	.60	-				
3.	ACT RHETORICAL PROFICIENCY-FORMAT II	.53	.64	-			
4.	ACT TOTAL (ENGLISH USAGE, RHETORICAL PROFICIENCY-FORMATS I AND II)	.89	.81	.81	-		
5.	SAT-VERBAL	.64	.66	.62	.75	-	
6.	SAT-TSWE	.74	.61	.50	.75	.76	-

The interrater reliabilities of the writing scores, as reported in Table 2, also suggest that there is interrater consistency in the scores assigned by raters using the holistic method as well as in the subscores assigned by raters in *language* and *organization*. The interrater reliability of the *audience* subscore, however, is low, which would suggest inconsistency in the ratings that constitute this score.

DISCUSSION

The premise that was investigated in this paper was that objective tests might possibly be useful as an indirect and "passive" method to assess the features or components of the composing process. The results of this study suggest that although there is a moderately high relationship between the objective measures and samples of actual writing, the skills measured in the two types of assessments (objective and writing samples) are not exactly the same. The suggestion holds especially true for the correlations between the objective measures and the *audience* writing subscore, where the relationships were low to moderate. Moreover, the skills assessed in the objective *usage* measures seem to be very similar--thus the relatively high correlations among the ACT English Usage, ACT Total, SAT-Verbal, and SAT-TSWE tests. Two objective measures, however, the ACT Rhetorical Proficiency-Formats I and II, correlated less well with the other objective measures, which suggests that these measures probably assess somewhat different skills than do the other objective measures. In addition, the data indicate that the rhetorical proficiency measures correlated less well with the actual writing samples. These results may have occurred because the measures of rhetorical proficiency

make far more complicated and sophisticated demands on reading skills than do the other measures.

Furthermore, the correlations between the objective measures and demonstrated writing proficiency are such that the highest correlations occurred between the *organization* and *language* writing subscores, on the one hand, and on the other, the ACT English Usage Test and the SAT TSWE, two tests that are usage-oriented in nature. This result concurs with the results of the study conducted by Godshalk, et. al., who, by way of explanation, suggested that "the characteristics of single sentences play an important part in the overall impression made on the reader" (p. 19). Since the skills assessed in the abbreviated version of the ACT English Usage Test heavily emphasized syntax, diction, and style--factors that also seem to dominate the rating criteria of the writing samples--the relatively high relationship between the two measures may thus be partly accounted for.

IMPLICATIONS

Collectively, these results suggest that at least some of the skills emphasized in the writing assessment seem to be measurable through objective assessment instruments. In particular, there seems to be a moderately high relationship between the holistic, *language*, and *organization* writing scores and the objective assessments. The *audience* writing score, however, not only seems to have a much lower correlation with the objective assessments, but also exhibits a lower interrater reliability than do the other writing scores. Further examination of this score should be made with particular attention given to the nature of the writing exercises and to the criteria actually used by raters in formulating this score for examinees.

The results of this study should be considered in light of the following limitations: 1) These data were based on three relatively small samples of students. 2) Only limited samples of writing could be obtained. 3) Because of time constraints, only abbreviated objective assessment instruments could be administered. Such limitations suggest that the external validity of the results is limited. However, since there is some evidence that a moderate-to-high relationship exists between these objective assessment instruments and demonstrated writing proficiency, such a relationship should continue to be the focus of intensive research. In fact, since much previous research indicates that objective usage tests seem to correlate well with demonstrated writing proficiency, research should continue to be directed toward resolving the following questions: (1) Are raters of writing assessments mainly influenced, however subliminally, by scribal errors? (2) Can the emerging research that pinpoints writing strategies, processes, and features in terms of specific rhetorical tasks be translated into the most judicious selection of such components for objective assessment?

If these questions can be resolved and the relationship between objective assessments and demonstrated writing proficiency proves to be high, then instructors could, for placement purposes, use such objective measures as gross indicators of writing ability. These measures would not only be more efficient in terms of administration time and money, but they would also be far easier to score fairly than are writing samples.

If, on the other hand, the objective measures do not prove to be highly related to writing proficiency, then the thinking/languaging processes involved in writing may be unique to the composing process and

thus not easily measurable in a comprehensive manner by objective test formats. Nevertheless, objective assessments may still provide *additional* information about students' rhetorical abilities that could be used in conjunction with writing samples. In such a case, objective assessments would provide useful information about students' rhetorical abilities that, in lieu of or in combination with information from actual writing samples, could be used with maximum efficiency and economy for placing students in rhetoric classes.

REFERENCES

Breland, H.M. A Study of College English Placement and the Test of Standard Written English. Princeton, New Jersey: Educational Testing Service, 1977.

College Entrance Examination Board. The Test of Standard Written English: A preliminary report. Princeton, New Jersey: Educational Testing Service, 1975.

Godshalk, F.I., Swineford, F., and Coffman, W.E. The Measurement of Writing Ability. New York: College Entrance Examination Board, 1966.

Steele, J.M. The assessment of writing proficiency via qualitative ratings of writing samples. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, 1979.